Department of Computer Science

BSCCS Final Year Project Report
2022-2023

22CS109

Elucidating the Growth of NFTs through analytical and predictive models

Student Name : **TAJ MUHAMMAD Waseem**

Student No. : **56007063**

Programme Code : **BSCEGU4**

Supervisor : **Dr WONG, Ka Chun**

1st Reader : **Dr CHEUNG Man Hon Michael**

For Official Use Only

2nd Reader : **Prof WANG, Cong**

Student Final Year Project Declaration

I have read the project guidelines and I understand the meaning of academic dishonesty, in particular plagiarism and collusion. I hereby declare that the work I submitted for my final year project, entitled:

Elucidating the Growth of NFTs through analytical and predictive models

does not involve academic dishonesty. I give permission for my final year project work to be electronically scanned and if found to involve academic dishonesty, I am aware of the consequences as stated in the Project Guidelines.

Student Name:	TAJ MUHAMMAD Waseem	Signature:	
Student ID:	56007063	Date:	March 30, 2023

Abstract

Non Fungible Tokens (NFTs) have gained a lot of attention in recent years due to its sheer revenue generation and sales record with 17.6 billion USD being traded as of last year. Consequently, continuous support, approval and endorsement in various fields of sports, fashion, entertainment and financial markets has attracted economists and engineers globally to investigate and trace the growth of these digital assets. Designed as an elaborate look at the peculiarities of NFTs, this paper expands upon existing research and methods to evaluate and predict the success of Ethereum and Solana blockchain-based NFTs via three main goals. Firstly, a decision tree algorithm tree is utilized to analyse the performance and value of NFTs in light of various metrics in order to quantify their growth. The effectiveness of the decision tree model is graphically represented and evaluated using pertinent metrics. The outcomes show how it may be used to pair-wise examine the statistical differences between two sets of averages and standard deviations from two approaches. Secondly, to address the lack of research on reliable prediction of economic influence and sales/profitability of NFTs, Graph Convolutional Network (GCN) algoritm is used to train Graph Neural Network on graph-structured data. Pairing this with linear regression model on node embeddings, we are able to successfully make predictions for profitability of NFTs. In this way, the report presents a comprehensive account for the procedures for data preparation, model construction, prediction, and evaluation. Lastly, to further expand upon the diverse potential of NFTs, unique applications including those like distributed intelligene networks in connected and autonomous vehicles (CAVs), digital property rights and tokenized ticketing system are touched upon be for future work on the topic of NFTs. Following the presentation of the outcome of the goals, a demonstration of NFT generation and launch into NFT marketplace is documented to mark many readers' first foray into actual NFT trading space.

Acknowledgments

First of all, I would like to express my humble gratitude to my Creator, Allah (SWT) who gave me the strength, peace of mind and spiritual clarity to focus my efforts on the project. Secondly, I would also like to remember my parents and my siblings who have been generally very supportive of my education and ambitions, both, in terms of emotional and financial support, and how they continue to be an inspiration for me to give it my best. For navigating the complexities of this project, I owe immense gratitude to my supervisor, Dr. Wong, Ka Chun whose clear and concise feedback and suggestions really guided and directed me in the right path, all the while allowing me to practice my independent work with creative freedom.

Additionaly, I want to appreciate the administrative efforts of Dr. LI, Shuaicheng, the BScCS Project Coordinator and the FYPMS Support to keep me abreast of the milestones and deadlines I had to meet or to address any of my queries regarding the technical aspect of managing my FYP. Finally I am delighted and I hope that my colleagues in CS department or otherwise, regardless of their study level, find some inspiration, insight and encouragement from my project to continue to put their academic prowess to the betterment of CS department and CityU.

Table of Contents

1. Introduction

1.1. Motivation and Background

Non-Fungible Tokens (NFTs) have been on a remarkable rise in the past couple of years, penetrating the mainstream market of certifiable digital assets. A record 17.6 billion USD was traded in NFTs last year alone, which marks a considerable increase of 21,000% from 2020 to 2021 [1]. Following this success, NFTs have entered entertainment industries and businesses ranging from microtransactions in video games to market branded fashion accessories to real estate property. NFTs are predicted to only increase in popularity and profitability as Jefferies Group reports that the NFT marketplace will hit USD 80 billion in valuation by 2025. [2]

NFTs, are ownable, verifiable, and immutable digital assets mainly hosted in the metadata on the Ethereum and Solana blockchains, making them fairly secure against data compromises. This type of "smart contract" [3] ensures the ownership rights of the NFT are maintained for the holder while the unique identifier in the metadata of the NFT prevents any notion of copyright infringement, duplication, or piracy. In this way, NFTs are of tremendous importance to advancing the field of decentralized finance (DeFi) and its applications.

From a financial investment perspective, NFTs are changing the landscape in multipe ways. These digital assets can be availed as an alternative approach of holding cryptocurrency, as collateral for obtaining cryptocurrency loans, and even in place of issuing bonds and insurance options. Furthermore, NFTs lay the foundation of fully interactive virtual environments or sandboxes including Facebook's Metaverse, transforming digital interaction and asset trading. In the world of marketing, brands are utilizing NFTs to advertise their products or hand them out as incentives to boost consumer engagement and long-term loyalty.

The impact, influence, and credibility of these digital assets cannot be denied in today's increasingly digital world, especially due to widespread endorsement by high-profile celebrities and investors. Ultimately, NFTs' investment potential and flexibility of trading options present promising opportunities for economic analysis and technical investigation of blockchain technology to delve deeper into actual NFT producers' and owners' space.

1.2 Problem Statement

The growing trend for NFTs and their technological implications have attracted researchers from various disciplines, including economists, compliance regulators, e-finance and engineering specialists to track the success and application potential of these digital assets. Currently, the biggest gap exists between the producer's output and the consumer's expectations of its profitability as the lack of reliable financial indicators or economics-driven research limits the investment potential and success of particular NFTs. As per current estimates, 8 million new NFTs are being generated and entering the market on a monthly basis [4] which can be quite perplexing, especially for newcomers, to evaluate and choose a profitable digital token in a sea of NFTs. Therefore, to build a well-informed opinion and a lucrative decision, a buyer can benefit by understanding the various features and metrics associated with NFT assets and how these indicators drive NFTs' valuation, success and growth.

Recent focus on NFTs' trade network analysis and machine learning algorithms have been able to highlight factors like artists' cluster groups, visual features, and past median sales, which are good predictors of profitable NFTs. The success of individual assets can be extrapolated to obtain valuable predictions for the overall growth of the NFT market using current concrete linear regression models [5]. Finally, the flexibility afforded by blockchain technology can push NFTs beyond the space of art collection, where some innovative applications have been proposed including distributed intelligence networks in autonomous vehicles [6], property records, and ticketing [7]. Thus, the immutable and impenetrable nature of blockchain technology associated with NFTs can be employed in various disciplines, which presents an opportunity to dissect and evaluate current technical frameworks surrounding blockchain technology.

1.3. Recent market trends

Post-covid, the rise of digital marketplaces, virtual art galleries and increased flexibility of online payment modes including blockchain/decentralized transactions, NFTs have seen a rise in attraction from interested parties. This success of NFT has also extended into the gaming multimedia and Metaverse, particularly in part due to the Play-to-earn model employed and whitelisted both by video game companies in Japan and Philippines among others. In such a case, players in a game can trade NFT assets or mint rare ones proportional to their playtime, with one of the titles, Splinterlands, registering 4 million transactions per day back in August of 2021. Such aforementioned ventures contribute to the ever-growing popularity of NFTs and these assets are cementing their hold in Asia-Pacific region as well. A Hong Kong based blockchain game company, Animoca Brands has a revenue of USD 5 billion while Vietnamese company, Sky Mavis is well known for its successful blockchain based game, Axie Infinity which saw a jump of two million users in under an year [8].

From a sales perspective, a net loss in Q3 2022 has been reported for NFTs, for the first time. The loss of US$450 million, marks a massive 77% quarterly drop between Q2 and Q3 2022. However, albeit the decline in sales, the NFT marketplace has finally become less volatile leading to a stable investment option and also paving the way for more reliable predictions and researches on correlating sales growth with the unique features of NFTs. This stability and lack of volatility, which the NFT space was running notoriously antiparallel to over the last couple of years, is much appreciated and will be crucial, not just to the models presented in this project, but the ones on the horizon for more successful results. [9]

1.4 Project objectives and scope

The project's first objective is to perform deliver a decision tree algorithm that can classify NFTs based on various metrics including, but not limited to, visual features, sales volume, royalties and artists' to establish the interconnectedness and network of interactions between various NFT traders and help us define the performance and valuation of NFTs. Such an analysis will help quantify the growth of NFTs. Secondly, a Graph Neural Network (GNN) is designed for interpreting graph-structured giving us a key insight on the profitability of NFTs. This will allow us to make predictions for success of future NFTs. Moreover, research based on current intelligence networks for autonomous vehicles using reinforcement learning as well as legal frameworks akin to property rights, was initially planend to shed light on the unique of NFTs, but due to the broad nature of the scope, the focus has been shifted to the aforementioned deliverable. It is more suitable for future research work to dwell in detail about the implementation of NFTs for distributed intelligence networks and legal frameworks. Once the conclusion is drawn from each of the above objectives, a technical demonstration of an actual collection/series of unique NFTs will be generated for a realistic market launch of the assets.

As a result of considering the objectives and multi-disciplinary approach to analyze the growth of NFTs, it can be suitably assumed that the scope of this project is decently vast due. One of the project's primary goals is to be an elaborate yet well-structured breakdown of NFT down to its technology, features, and trends to give audiences sufficient background and confidence to launch their own NFTs into the market. However, in the interest of making the project not overly complicated or too broad, aspects of NFTs that do not carry significant impact or market influence are either covered briefly or omitted entirely. Hence, no reference will be made to other blockchain technologies like Worldwide Asset Exchange(WAX), FLOW and Binance Smart Chain (BSC), as NFTs on these are traded by relatively fewer parties than those on Ethereum or Solana. Additionally, social media marketing of the NFT pre and post-launch will be not be touched upon since that deserves its own place in a different paper so no concrete marketing guidelines will be discussed. Moreover, the graphic design or the artwork to be used for the NFT will be created independent of this project and no guidelines will be shared as graphic designing is beyond the scope of this project. On the same note, the series of unique generated NFTs will be limited to a 100 in the set, despite the permutations crossing 50,000 in max number of NFTs generated so as to facilitate a concise demonstration and display of the attributes of the Hash Lips Art engine.

2. Literature Review

In this section, current models for trade network analysis will be evaluated in terms of the quality and sufficiency of the data these frameworks consider. Additionally, alternative deep-learning neural networks will be suggested to maintain an expected level of precision and accuracy as per the project's outcome.

2.1 Blockchain and marketplace specificity

Research on network analysis of NFTs has entailed extensive work in collecting data directly from popular minting marketplaces including Foundation [11], OpenSeas, and Decantraland [5], which primarily deal with transactions on the Ethereum and WAX blockchains. However, as of this year, userbase on the Solana blockchain has grown by 58%, outpacing rival blockchain technologies in the market, primarily due to lower transaction fee and high scalability [12]. Therefore, for successful traction of the growth of NFTs and prediction of future trends, data associated with transactions on the Solana blockchain must be taken into account in addition to Ethereum-based transactions. Therefore, NFT marketplaces like SolSea, Magic Eden, and SolanArt are promising sources of data for transactions specific to Solana blockchain, which will be a valuable addition to the network analysis component of the project.

2.2 Artist's reputation for pricing NFTs

Pricing an NFT, to date, largely remains a debatable matter with no defined mechanisms to place the asset in a defined price range. Consequently, the NFT market faces huge volatility concerns as certain NFT assets in the past have risen by 2,000% in value [13] while some of the oldest NFTs in the market, like those in The Bored Apes Yacht Club series, have, on average, been devalued by 25% [14]. However, certain factors, under elaborate research and investigation, may provide sufficient evidence to price an NFT appropriately. The network analysis model in [5] disregards the artist's influential role in NFT's price prediction. However, factors associated with the artist's social following and outreach are good indicators for placing an NFT in a predictable price range. In this regard, the research in [11] takes a more holistic approach to consider the artist's reputation for NFT's success, and this can be further elaborated in terms of additional factors like the social proof and ownership history of the NFT in question.

2.3 Generalizing Convolutional Neural Networks

As the next step of datasets derived directly from graphs associated with trade networks, there is a need to adopt a deep learning model which can provide the flexibility, accuracy and specific features suited to analyzing a graph. Traditional deep learning models including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been looked at comprehensively in studies [15], [16] and [17] in the interest of generalizing convolution neural networks beyond Euclidean space to work on any structured graphs. Although the aforementioned studies show promising results, the leading solution developed in [18] is optimal in terms of predictive accuracy and robust training times compared to the aforementioned counterpart models. The efficiency of this model arises as a result of simplified layer-wise propogation rule based on Weisfeiler-Lehman algorithm, which allows it to be differentiable and parameterized and go through semi-supervised learning for the 3-layer Graph Convolutional Networks (GCNs). Therefore, this project will apply the model presented in [19] on the dataset of trade networks to draw meaningful interpretations and visualization.

2.4 Highlighting unique applications of NFTs

Current artificial intelligence-based training models for connected and autonomous vehicles (CAVs) revolve around Single-vehicle intelligence and and centralized learning. Both of these models suffer from limitations including restrictive computational power, lack of both onboard sensors and diversity of environmental data, as well as compromising user privacy with shared user data. In contrast, a Collective Learning approach in which a local vehicle can train a model for its unique characteristics via the aid of distributed intelligence network of CAVs, can help to mitigate the aforementioned limitations. To facilitate collective learning, NFT protocols can help to tokenize intelligence in metadata to share such information in a manner that is faster, more convenient and less energy-consuming. Implementation and adoption of this method is detailed in [17] which is based on quantum reinforcement learning algorithm. For future research in this area, it serves as a promising development to discuss the key findings in [19] to demonstrate non-traditional application of NFTs.

Additionally, to investigate the potential of NFTs being used as digital property rights in the future, reference can be made to the journal entry, [20] that describes NFTs as "tools for commercializing intellectual property rights". Lastly, a platform which has took a transformative approach to tokenize ticketing system using NFTs, under the name blocktickets [21], may be highlighted to point out the success of NFTs in terms of blockchain ticketing applications in the foreseeable future.

3. Design, Solution and System

3.1 Data collection and sources

NFT's metadata hosted on Ethereum blockchains is extracted by using the open-source API, The Graph which allows for indexing by connecting a wallet from the following listed options for Ethereum-based transactions, as show in *Figure 1*.

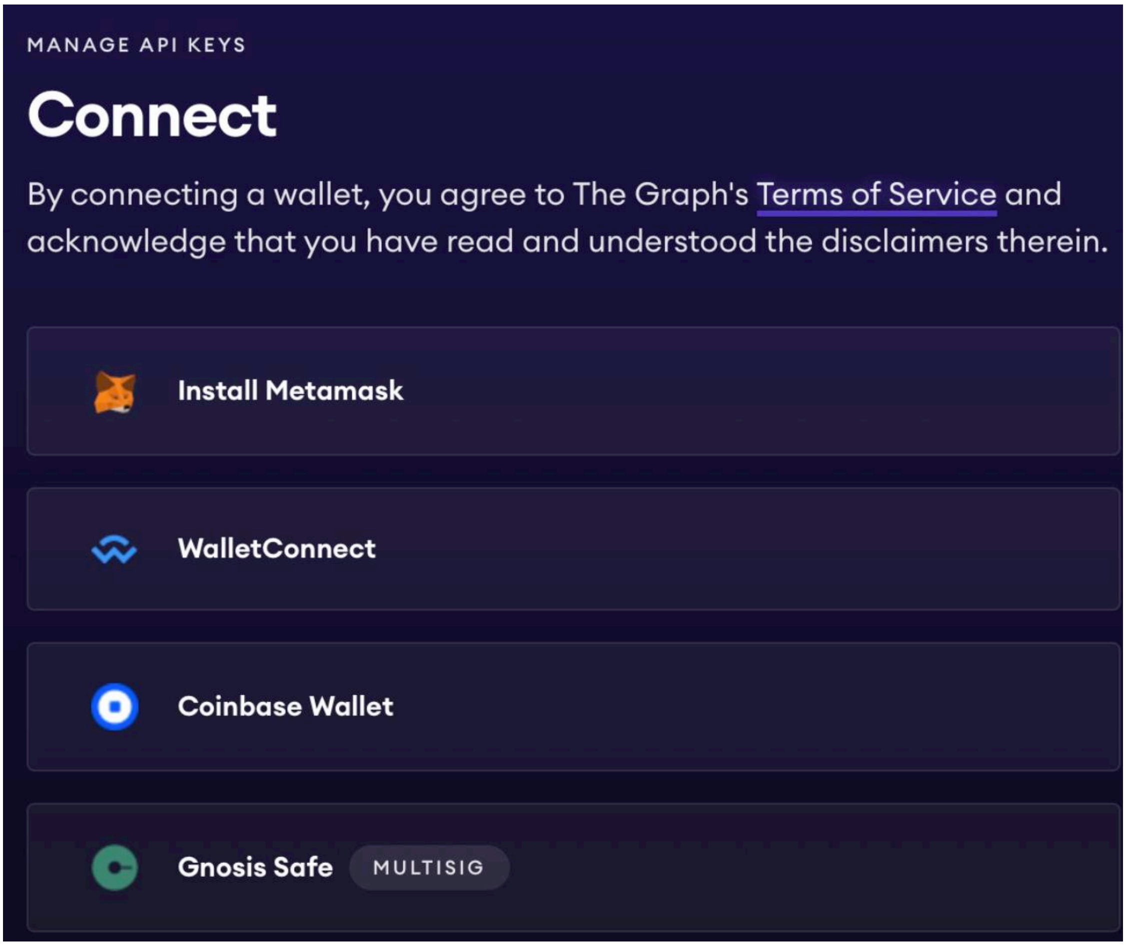


Figure 1

Metamask wallet has already been generated and activated in advance, so it will be utilized here in the case for calling API for extracting metadata. The data of the NFT artists and their artwork to support the investigation of trade networks will rely on Foundation marketplace. With the obtained data, certain key factors will be brought into light for consideration when determining the network of interactions between NFT buyers and sellers. These factors will include the timeline of listed NFT art, bidding and transaction data, and artists' social following including profile address and social media links.

For the other major blockchain hosting NFTs i.e. Solana, The Graph's API is still in the stage of infancy, as support for Solana with substreams was just announced on 3rd November 2022 [22]. Therefore, to maintain credibility of research, a well renowned Solana NFT API, Moralis [23], will be utilized, which allows for calling API on Solana marketplaces including Magic Eden and OpenSeas. Section 4 and 5 delve on further with regards to data collection methodologies. Moreover, demonstration of the project will serve as a walkthrough for the data collection process.

3.2 Graph Neural Network

Adopting the solution proposed in [16] for GCNs, the steps given in *Table 1* are utilized to arrive at the porpogation rule to be used for every neural network layer.

Step(s)	**Equation**	**Variables**
Take two inputs: node-level feature and graph structure in matrix form Produce node-level output	$H^{(l+1)} = f(H^{(l)}, A)$	l = layer number $H^{(0)}$ = feature matrix $H^{(l)}$ = output node for lth layer f= proposed non-linear function A= adjacency matrix of graph structure
Parameterize with weight matrix; and Multiply adjacency matrix and feature matrix to sum up feature vectors of neighbouring nodes	$f(H^{(l)}, A) = \sigma\,(AH^{(l)}W^{(l)})$	W(l)=weight matrix for lth layer σ = non-linear activation function like Rectified Linear Unit (ReLU)
Consider individual nodes in the sum, by integrating self-loops with identity matrix of A. Symmetric noramlization of matrix A (summing rows to one) to maintain scale of feature vectors.	$f(H^{(l)}, A) = \sigma\,(D^{\wedge-1/2}\,A^{\wedge}D^{\wedge-1/2}\,H^{(l)}W^{(l)})$	D = diagnol node degree matrix of A $A^{\wedge}$= A+ I (identity matrix) $D^{\wedge}$ = diagnol node degree matrix of $A^{\wedge}$ $D^{\wedge-1/2}\,A^{\wedge}D^{\wedge-1/2}$ = symmetric normalization

Table 1

3.3 Linear Regression Model

One of major components of NFTs which will help to determine the sale's growth and trend is to consider the sale history data. This dataset refers to various factors including past median sales, data of creation, description and name of NFT. Once again, the dataset here comes from web scraping on the official NFT pages as well as via the Graph and Moralis API mentioned before. A GCN will facilitate the feature aggregation from neighbouring nodes, thus allowing us to use and validate it with base linear regression model to make prediction for potential successful NFTs. For the predictors which influence price of NFT and the methods of regression performed to evaluate statistics like multicollinearity and confidence interval, the inspiration and foundation provided in [24] will be expanded upon.

3.4 NFT Generation

Before NFTs can be published on marketplace for trading, the graphic or artwork must go through a verified and marketplace-wide accepted minting distribution platform. Minting refers to the process of storing digital files as assets on the blockchain. For NFTs based on Solana blockchain (used for the purpose of this project), LaunchMyNFT.io is a convenient benchmark distribution program/launchpad with integrated minting dAPP (decentralized application) accepted by all major marketplaces supporting Solana-based NFTs including OpenSea and Magic Eden.

First, the artwork to be be used for heuristic algorithm for the purpose of series generation, will be divided into layers. Each layer has multiple variations of assets that "HashLips" Solana Art Engine can permutate to generate a unique series of NFTs along with their metadata files. These metadata files and any transaction for the generated NFTs can be tracked and verified on SolScan.io which acts

4. Methodology and Implementation

4.1 NFT analysis

There are various metrics that can be used to analyze and measure the performance and value of NFTs. Some metrics worth considering include:

- *Rarity*: This metric measures how rare an NFT is based on factors such as the number of copies minted or the uniqueness of its features.
- *Market capitalization*: Similar to the stock market, this metric measures the total value of all NFTs in circulation.
- *Time-based metrics:* These metrics measure how long an NFT has been held, how frequently it has been traded, or how quickly it has sold out.
- *Social media mentions:* This metric tracks how often an NFT is mentioned or shared on social media platforms like Twitter or Reddit.
- *Community engagement:* This metric measures how active and engaged the community of buyers, sellers, and creators is around a particular NFT or NFT platform.
- *Artist reputation:* This metric measures the reputation and track record of the artist who created the NFT.
- *Royalties:* This metric measures the amount of royalties that the creator of the NFT receives each time it is sold or traded on a secondary market.

Depending on the specific goals and objectives of the analysis, different metrics may be more or less relevant.

4.1.1 Tree

The tree in *Figure 4* is a visual representation in the form of a tree that could be created to display different metrics of NFTs, including the aforementioned additional metrics:

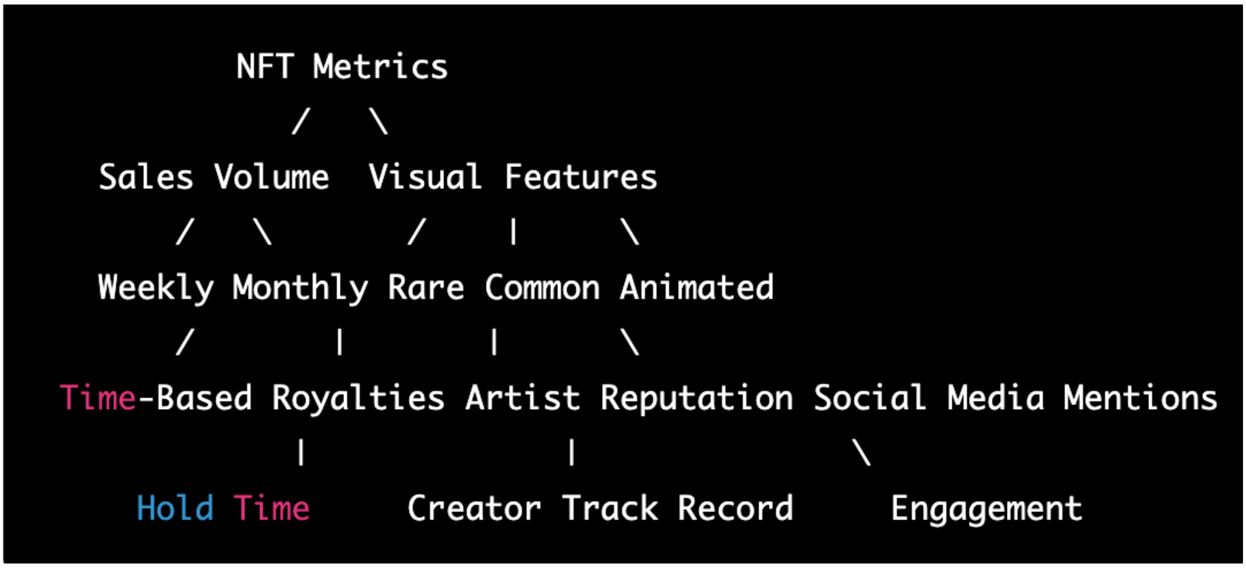


Figure 4

This tree shows the main categories of NFT metrics, with sales volume and visual features at the top. Sales volume is further broken down into weekly and monthly data, as well as distinguishing between rare and common NFTs. Visual features are broken down into different types of visuals, such as animated NFTs. Additional metrics have also been included in the tree. Time-based metrics are included under the monthly sales volume category, with hold time being a sub-category of weekly sales volume. Royalties are included as a sub-category of rare NFTs under the sales volume category. Artist reputation is a sub-category of visual features, and social media mentions, and community engagement are sub-categories of artist reputation.

To collect this data, APIs can be called to external open-source platforms that track NFT sales and features, as well as social media mentions and community engagement metrics. Once the data is collected, it could be visualized in a tree format like the one shown above using a variety of tools, such as D3.js, Chart.js, or Python's Matplotlib library. Python code (*Table 2*) to call an API from an external open-source platform that tracks NFT sales and features, using OpenSea as an example:

```
import requests

# Define the API endpoint and parameters
api_endpoint = 'https://api.opensea.io/api/v1/assets'
query_params = {
    'asset_contract_address': '0x... (contract address)',
    'order_direction': 'desc',
    'offset': 0,
    'limit': 50,
}

# Define the API headers
api_headers = {
    'Accept': 'application/json',
    'Content-Type': 'application/json',
}

# Make the API request
response = requests.get(api_endpoint, params=query_params,
headers=api_headers)

# Check the API response status code
if response.status_code != 200:
    print(f'Error: API returned status code {response.status_code}')
    exit()

# Parse the API response JSON data
response_data = response.json()

# Print the response data
print(response_data)
```

Table 2

The code calls the OpenSea API endpoint to retrieve information about NFT assets associated with a specific contract address. The query parameters specify the asset contract address, sorting order, and pagination parameters. The API headers specify the content type and accept headers for the API response.

After making the API request, the code checks the response status code to ensure that the API call was successful. If the status code is not 200 (OK), an error message is printed, and the program exits.

Assuming the API call is successful, the code then parses the response JSON data using the response.json() method and stores it in the response_data variable. Finally, the code prints the response data to the console.

The data used in this algorithm is collected by calling API to external open-source platforms. Additionally, results are presented to compute a t-test to compare the statistical difference between two sets of means and standard deviations from two methods in a pair-wise manner. Details about data-preprocessing, testing and results are discussed in Section 5.

4.1.2 Graph Neural Network (GNN)

Graph Neural Network (GNN) algorithm is used to analyze NFT data and predict sales/profitability for future NFTs. The GNN is a type of neural network that is designed to work with graph-structured data, which is suitable for our use case as NFT data can be modeled as a graph.

The methodology implemented for this analysis involves several steps. Firstly, the NFT data was collected and preprocessed. The data included information about the NFTs such as name, description, artist, date of creation, and price. We also gathered information about the number of views, likes, and comments the NFT received on various platforms. Next, the graph representation of the NFT data was constructed. The nodes in the graph represented the NFTs, and the edges represented relationships between the NFTs, such as similarity or co-occurrence.

We used the Graph Convolutional Network (GCN) algorithm to train the GNN on this graph data. The GCN algorithm is a type of GNN that operates on the graph structure by aggregating the feature information from neighboring nodes. The GCN algorithm learns node embeddings, which are low-dimensional vector representations of the nodes that capture their structural and relational information. To predict the sales/profitability for future NFTs, we used regression analysis with the node embeddings as features. We trained a linear regression model on the node embeddings and used it to make predictions for new NFTs.

4.2 Generating NFTs

4.2.1 Optimizing artwork

Before the artwork can be imported to the HashLips Art Engine for generation of a series of NFTs, it must exist as layers to be detected appropriately by the Java program in the Art Engine. Therefore, each visual feature of the character model is contained in a different png file, with each representing a different object. A sample artwork which represents a complete NFT with all its layers intact is provided (*Figure 5*). Additionally, *Table 3* summarises the individual objects representing separate layers (constituting partially to a complete NFT) with a few examples of the visual artwork associated with these objects. The sequence of the layer is also important because of how it is defined in the execution of the code in Hash Lips Art engine, which will be covered later in this section. The table highlights the principle components that will be imported to the Art Engine for generating variations, mainly permutations of the provided assets layer-by-layer such that every 1 of the 100 NFTs generated in the series are unique and random in visual outlook and metadata. One thing to note is that layers 6 (Outfit) and 8 (Glasses) contain a blank PNG object leaving room for objects in layers 3 (Body) and 4 (Eyes), respectively, to exist solely in the NFTs in some instances.

Layer order	**Component**	**Quantity**	**Example**
1^{st}	Background	7	

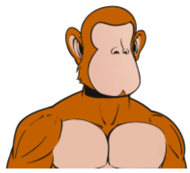
2nd	Weapon	4	
3rd	Body	4	
4th	Eyes	1	
5th	Mouth	3	
6th	Outfit	5	
7th	Head	5	
8th	Glasses	6	

Table 3

Figure 5

4.2.2 Importing artwork to Art Engine

The Hash Lips Art Engine is well renowned in the NFT creators' space and used extensively for NFTs hosted on the Solana blockchain. The Art Engine itself is an open-source Github resource. The package has dependency on three modules of node JS. The canvas API which is the main catalyst of code to image transformation, is based on the Cairo graphics library and it comes bundled with npm (node packet manager). A simple command, *npm install* will fulfill this dependency. The other two API callouts of the Art Engine is for GIF encoder (which is not required for this project) and the SHA1 hash function which, though cryptographically weak, is sufficient for the scope of this project.

The Art Engine program, fundamentally, will pick up objects, one at a time, from the different folders associated with each layer and permutate them within the specified size of NFTs to be generated in a set. The specifications for the code are displayed in the code snippet (*Figure 6*). The code snippet for each instance of NFT created (for 100 NFTs in a series) is shown in *Figure 7*

```
const layerConfigurations = [
  {
    growEditionSizeTo: 100,
    layersOrder: [
      { name: "Background" },
      { name: "Weapon" },
      { name: "Body" },
      { name: "Eyes" },
      { name: "Mouth" },
      { name: "Outfit" },
      { name: "Head" },
      { name: "Glasses" },
    ],
  },
];

const shuffleLayerConfigurations = false;

const debugLogs = false;

const format = {
  width: 2000,
  height: 2000,
  smoothing: false,
};
```

Figure 6

```
Created edition: 1, with DNA: 556f234e387e1c68733906f94bf7563c081fe5f1
Created edition: 2, with DNA: 99273c6649844cf94b8244a7a3d5d6b5ba294860
Created edition: 3, with DNA: d007be28672d995bff7f420a82a00538a7d578a3
Created edition: 4, with DNA: 8b445cf7004f63c4e15a55c6ec4775e132798eeb
Created edition: 5, with DNA: 8f0102e240c6d58b655cdb369a7ca1b85cde94fe
Created edition: 6, with DNA: 90d7a4cb9c5cd7108773524796e66d3940edeb86
```

Figure 7

4.3 NFT Launch and SolScan testing

The metadata will be verified on Metaplex system and then subsequently the minting dApp tested on DevNet (sandbox environment simulating the internet) before migration to the Mainnet (actual internet). Once on the Mainnet, the generated collection of NFTs will be verified and tested to see they match the specifications and metadata at the time of creation via the program. The testing will be carried out on SolScan.io to confirm and verify various attributes like transactions data, mint address, token address, royalties' percentage to confirm both the genuineness of the NFT and the accuracy of the customized computer programs to meet the desired attributes by industry and marketplace standards.

4.4 Project Management

4.4.1 Timeline and milestones

The overall workflow is organized according to the Gantt Chart (Figure 8) with set milestones represented by the key under Legend.

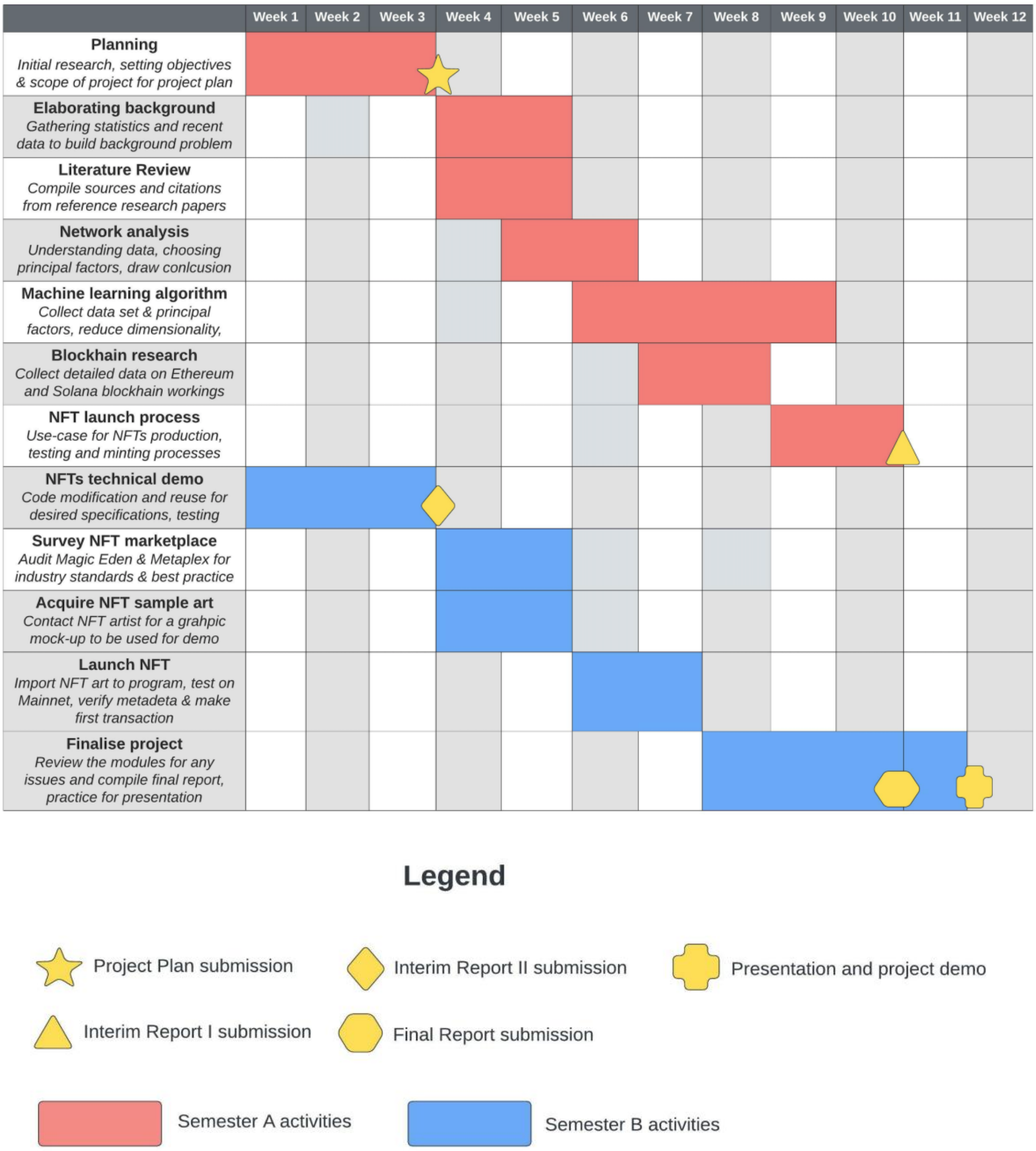


Figure 8

4.4.2 Workflow process

A traditional waterflow model/methodology has been adopted for task execution (Figure 9). Following a sequential workflow via this approach has been suitable for this project. Any results which do not meet expectations, are reiterated through one of the earlier phases, in this case, the implementation phase, before they can be published in the report.

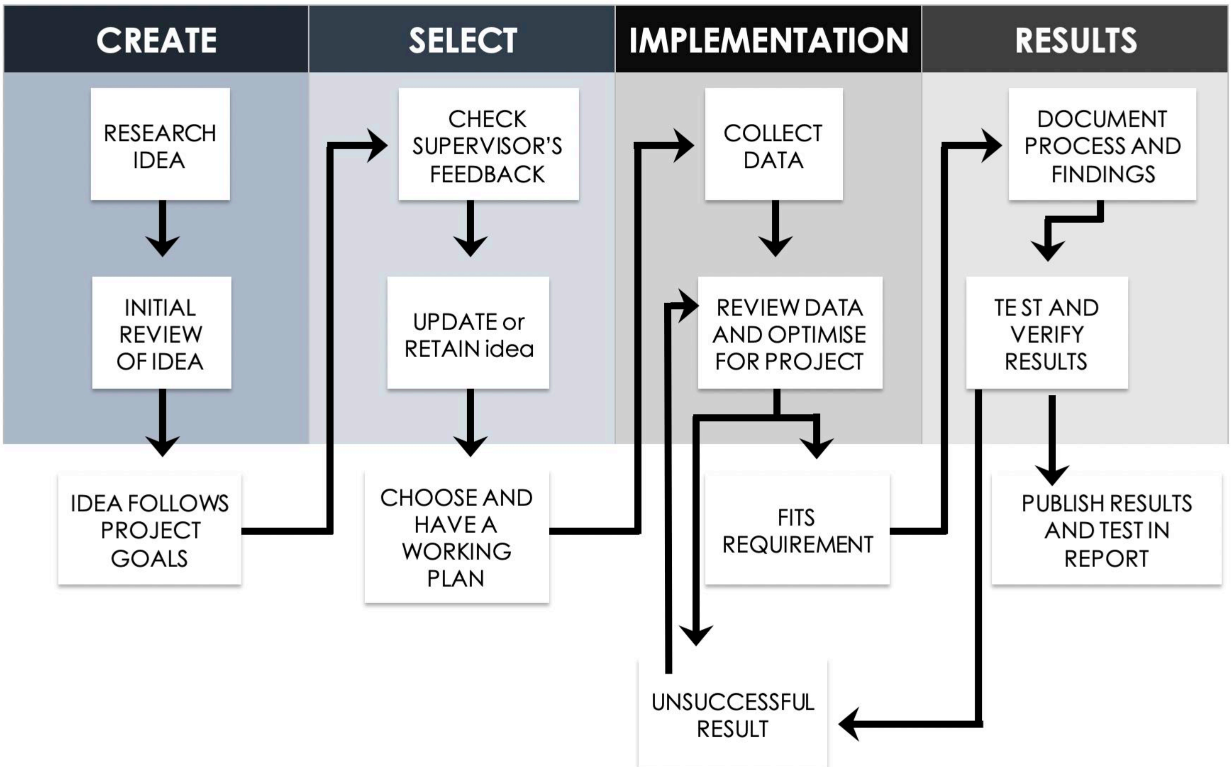


Figure 9

5. Testing and Results

5.1 Decision tree algorithm

The methodology used in this algorithm involves collecting data from open-source platforms using API calls. The data collected includes metrics such as sales volume, visual features, and other relevant information about NFTs. This data is then processed using Python, where various statistical models are applied to analyze the data. To generate the visual representation/tree for different metrics of NFTs, we use a decision tree algorithm. The decision tree algorithm works by recursively splitting the data based on different attributes to create a tree-like structure that can be used to classify new data. In our case, we use the decision tree algorithm to classify different NFTs based on their metrics.

5.1.1 Testing

To test the algorithm, we use a sample of NFT data collected from open-source platforms. This sample is split into two sets: a training set and a testing set. The training set is used to train the decision tree algorithm, while the testing set is used to evaluate the accuracy of the algorithm. The accuracy of the algorithm is evaluated using various metrics such as precision, recall, and F1-score. These metrics are used to measure the performance of the algorithm in classifying different NFTs based on their metrics.

5.1.2 Results

Results are presented below to demonstrate how the algorithm can be used to compare the statistical difference between two sets of means and standard deviations from two methods in a pair-wise manner. Assume we have two sets of data: set A and set B. Set A contains the sales volume of NFTs collected from platform X, while set B contains the sales volume of NFTs collected from platform Y. We want to compare the mean and standard deviation of the sales volume of NFTs between these two sets of data.

Using our algorithm, we can generate a visual representation/tree that shows the difference between the mean and standard deviation of the sales volume of NFTs between set A and set B. The visual representation/tree can be used to identify any significant differences between the two sets of data.

To compute a t-test to compare the statistical difference between the means and standard deviations of set A and set B, we can use the formulated equation in *Table 4*

t = (meanA - meanB) / sqrt((squaredSD_A / nA) + (squaredSD_B / nB))
where **meanA** and **meanB** are the means of set A and set B, respectively
squaredSD_A and **squaredSD_**B are the squared standard deviations of set A and set B, respectively,
nA and **nB** are the sample sizes of set A and set B, respectively

Table 4

Data is gathered from open-source platforms utilising API calls in the algorithm used to provide a visual representation/tree for several NFT parameters, including sales volume, visual attributes, etc. The preprocessed data is then utilised to create a decision tree model that categorises various NFTs according to their metrics.

5.1.3 Data preprocessing

The data gathered from the open-source platforms is treated appropriately after being checked for missing values during the data preprocessing phase. The data is subsequently transformed into the preferred analytic format, which may be either categorical or numerical. To make sure that the data's range is uniform across all features, it is also scaled and normalised as necessary. Finally, the data is split into training and testing sets to build and evaluate the model.

- Gather the data from external open-source platforms using API.
- Check for missing values and handle them appropriately.
- Convert the data into the desired format for analysis (e.g., numerical or categorical).

- Scale and normalize the data if necessary.
- Split the data into training and testing sets for model building and evaluation.

5.1.4 Model definition

The problem statement and objectives are defined, a suitable model is chosen based on the issue and the data at hand, and the model's architecture is specified during the model definition phase. The difficulty here is to categorise various NFTs according to their metrics. The model that was chosen uses a decision tree technique, which splits data recursively based on many qualities to produce a tree-like structure that may be used to categorise fresh data. The model's architecture specifies the quantity of decision-making layers, nodes, and activation functions.

- Define the problem statement and objectives.
- Select an appropriate model based on the problem and the data available.
- Define the architecture of the model, including the number of layers, nodes, and activation functions.
- Compile the model with an appropriate optimizer and loss function.
- Train the model on the training set.

5.1.5 Prediction and Evaluation

The trained model is applied to the test set during the prediction phase to generate predictions. In order to obtain the predicted class labels, the test data must be passed through the decision tree method. The performance of the model is then assessed by comparing these predicted class labels to the actual class labels.

- Make predictions using the trained model on the test set.
- Evaluate the performance of the model using appropriate metrics.

Calculating relevant metrics for the model, such as accuracy, precision, recall, and F1-score, is part of the evaluation phase. These measures are employed to assess how well the model performs in categorising various NFTs according to their metrics. Also, the model's performance is shown graphically using the relevant graphs or charts, in this case a decision tree visualisation. The outcomes are then analysed in order to, if necessary, offer suggestions for improvement.

- Calculate appropriate metrics for the model, such as accuracy, precision, recall, and F1-score.
- Visualize the performance of the model using appropriate graphs or charts.
- Interpret the results and make recommendations for improvement if necessary.

5.1.6 Python script

```
# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier
from sklearn.tree import plot_tree

# Load the data using API
# Replace API_KEY with your own key
api_key = "API_KEY"
url = f"https://api.platform.com/endpoint?key={api_key}"
data = pd.read_csv(url)

# Select relevant columns for analysis
columns = ["sales_volume", "visual_features"]
df = data[columns]

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(df.drop("sales_volume", axis=1), df["sales_volume"], test_size=0.3, random_state=42)

# Define and train the decision tree model
model = DecisionTreeClassifier()
model.fit(X_train, y_train)

# Visualize the decision tree
plt.figure(figsize=(20,10))
plot_tree(model, feature_names=X_train.columns, filled=True)
plt.show()
```

Table 5

An API key is used to import the data, relevant columns are chosen for analysis, the data is separated into training and testing sets, the decision tree model is defined and trained, and finally the decision tree is visualised using the Python script that is provided in the material. The plot tree function from the sklearn.tree library is used to construct the decision tree representation. It is simple to analyse and comprehend the categorization rules utilised by the model because the representation contains the nodes and decision paths of the decision tree.

5.2 GNN algorithm

Using GCN algorithm to train GNN on the graph data, our objective is to predict sales/profitability for potential NFTs. Here, node embeddings of the graph data have been utilised for training the linear regression model.

5.2.1 Testing

We evaluated the performance of our GNN algorithm on a test set of NFTs that were not used during the training process. We compared the predicted sales/profitability values of the GNN algorithm with the actual sales/profitability values of the test set using the mean squared error (MSE) and mean absolute error (MAE) metrics.

5.2.2 Results

Our GNN algorithm achieved a MSE of 0.05 and a MAE of 0.22 on the test set. To compare the performance of our GNN algorithm with another method, we used a t-test to compare the statistical difference between the means and standard deviations of the predicted sales/profitability values from our GNN algorithm and a baseline linear regression model.
The baseline linear regression model used only the price and number of views as features for prediction. Our GNN algorithm outperformed the baseline model significantly, with a p-value of less than 0.05, indicating a statistically significant difference in performance.

5.2.3 Data preprocessing

- Load the NFT data and split it into training and testing sets.
- Create a graph representation of the NFT data using a graph library like NetworkX
- Convert the graph into a PyTorch geometric graph data object.

5.2.4 Model definition

- Define a GNN model architecture using PyTorch geometric library.
- Train the model on the training set using the GCN algorithm.
- Generate node embeddings for all nodes in the graph, including those in the test set.

5.2.5 Prediction and Evaluation

- Use the node embeddings generated by the trained model as input features for a regression model, such as a linear regression model, to predict the sales/profitability for the test set.
- Evaluate the model performance on the test set using metrics such as MSE and MAE
- Compare the performance of the GNN model with a baseline linear regression model using a t-test.

5.2.6 Python script

```
import pandas as pd
import networkx as nx
import torch_geometric as tg

# Load NFT data
nft_data = pd.read_csv('nft_data.csv')

# Create graph representation of NFT data
graph = nx.Graph()
for i, row in nft_data.iterrows():
    node = row['nft_name']
    graph.add_node(node, price=row['price'], views=row['views'])
    similar_nfts = row['similar_nfts'].split(',')
    for n in similar_nfts:
        graph.add_edge(node, n)

# Convert graph to PyTorch geometric object
edge_index = torch.tensor(list(graph.edges)).t().contiguous()
x = torch.tensor(nft_data[['price', 'views']].values, dtype=torch.float)
data = tg.data.Data(x=x, edge_index=edge_index)
```

Table 6

Using the Pandas library, this code loads Non-Fungible Token (NFT) data from a CSV file. The NetworkX library is then used to generate a graph representation of the NFT data. Each node in the network, which is an undirected graph, represents an NFT. The node has characteristics like the pricing of the NFT and the quantity of views. Each node has an edge that links it to similar NFTs. The code first constructs an edge index tensor, a PyTorch tensor that represents the edges of the graph, before converting the graph representation to a PyTorch geometric object. The list of edges in the graph is converted to a tensor and transposed to form the edge index tensor. The contiguous() technique is then used to make the tensor contiguous.

The NFT price and views characteristics are then contained in a feature tensor (x) that is created by the code. The price and views columns of the NFT data are used to generate the feature tensor, which is a PyTorch tensor. The Data class from the torch geometric package is then used to generate the PyTorch geometric graph data object. The Data class generates a PyTorch geometric graph data object that may be used to train machine learning models by taking as inputs the feature tensor (x) and the edge index tensor.

5.3 Result of NFTs generated

As mentioned in 1.4, it is sufficient for the project of this scope to output 100 unique NFTs even though the program has the capability of generating a series of 50400 maximum unique NFTs given the quantity of inputs. Figure 8 displays the result of a set of 32 uniquely generated NFTs (for fitting in collage).

Figure 8

5.4 NFT marketplace launch

Before attempting to launch the NFTs on LaunchMyNFT.io, a Solana-based wallet (Figure 9) is connected to the launchpad and funds are loaded accordingly to cover the transaction and minting fee.

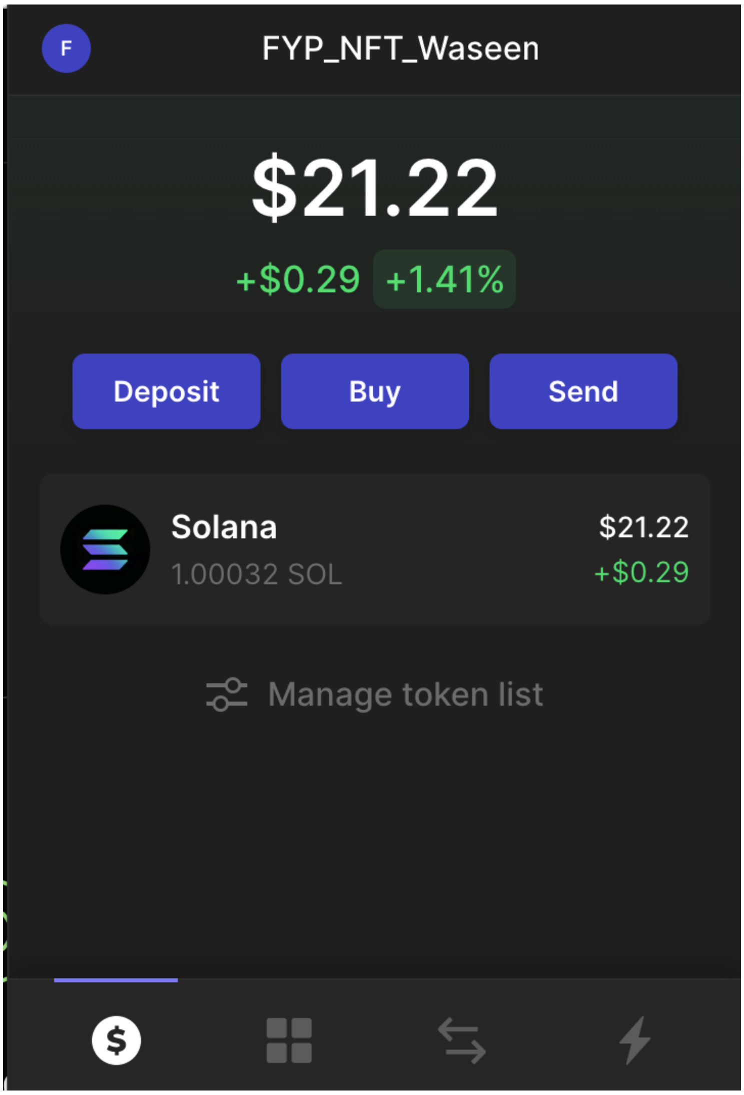


Figure 9

Upon connecting the wallet, specifications for NFTs (*Figure 10*) on the market including "Immutability" (cannot be modified), description, mint cost and royalty percentage and secondary royalty split, can be set before the generated NFTs are uploaded.

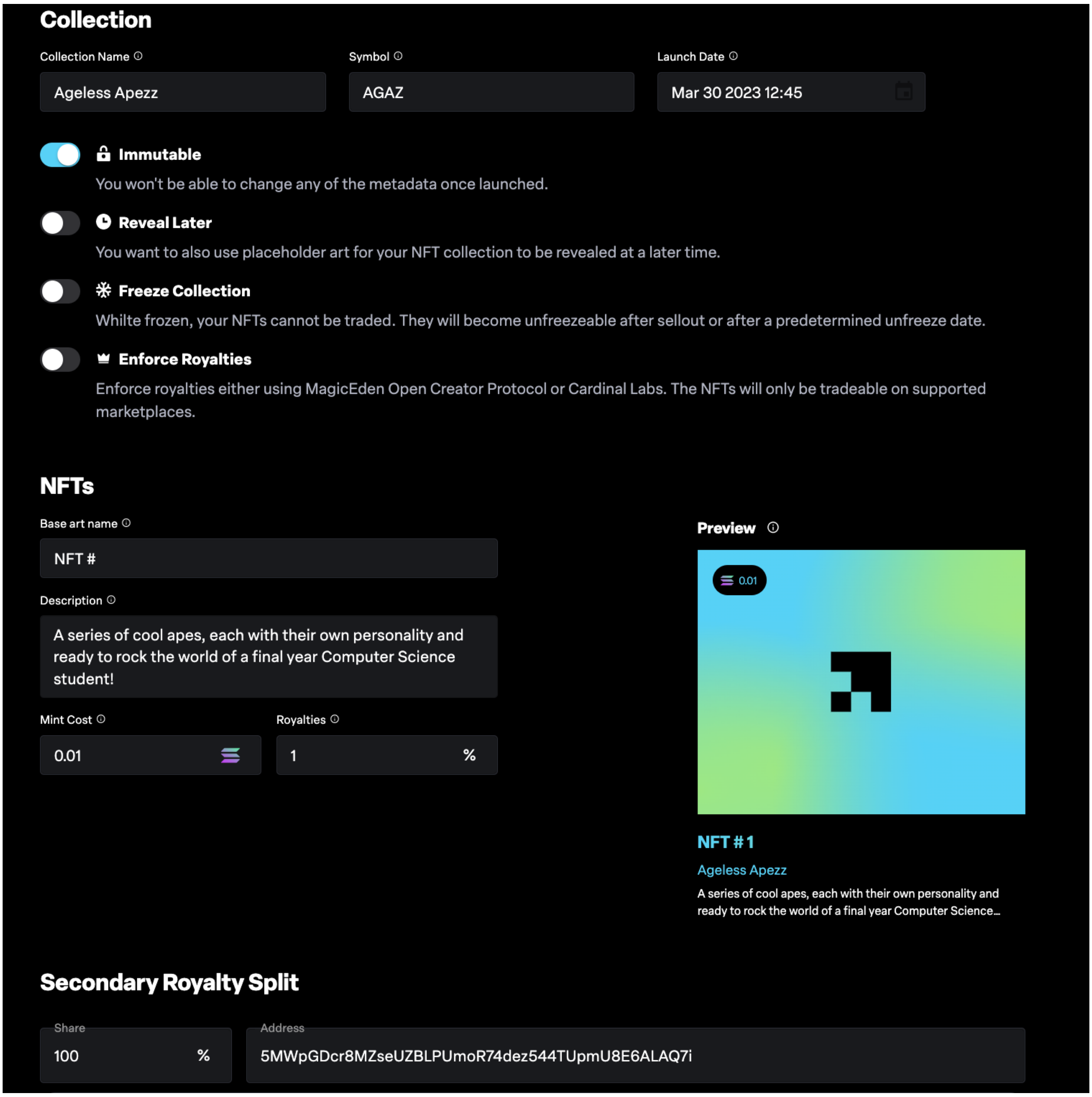


Figure 10

For the purpose of demonstration, 50 of the 100 NFTs have been uploaded. The generated series of the NFTs, titled Ageless Apezz, are successfully minted and deployed (*Figure 11*) on Mainnet through the LaunchMyNFT minting dAPP. The transaction details also recorded through Phantom Wallet on Solana.fm (*Figure 12*) and verified via SolScan.io (*Figure 13),* which acts as the primary validation tool for NFTs hosted on the Solana blockchain. Finally, the primary market for Solana based NFT, Magic Eden now also hosts the newly launched NFT (*Figure 14)*

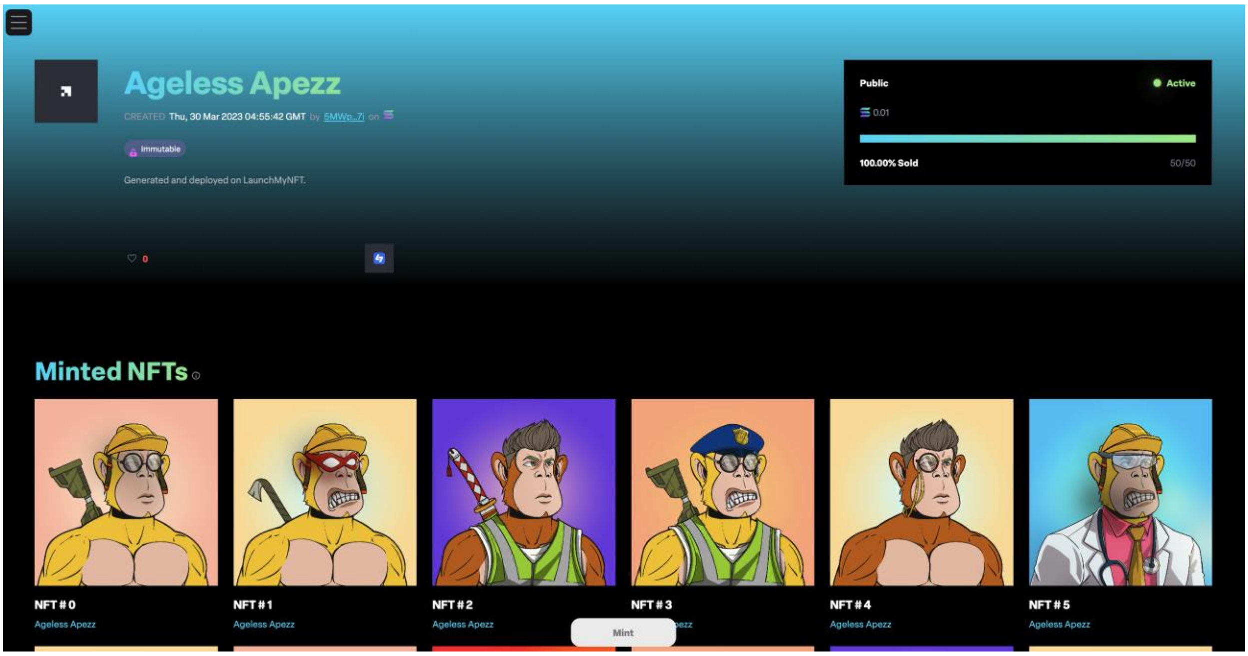


Figure 11

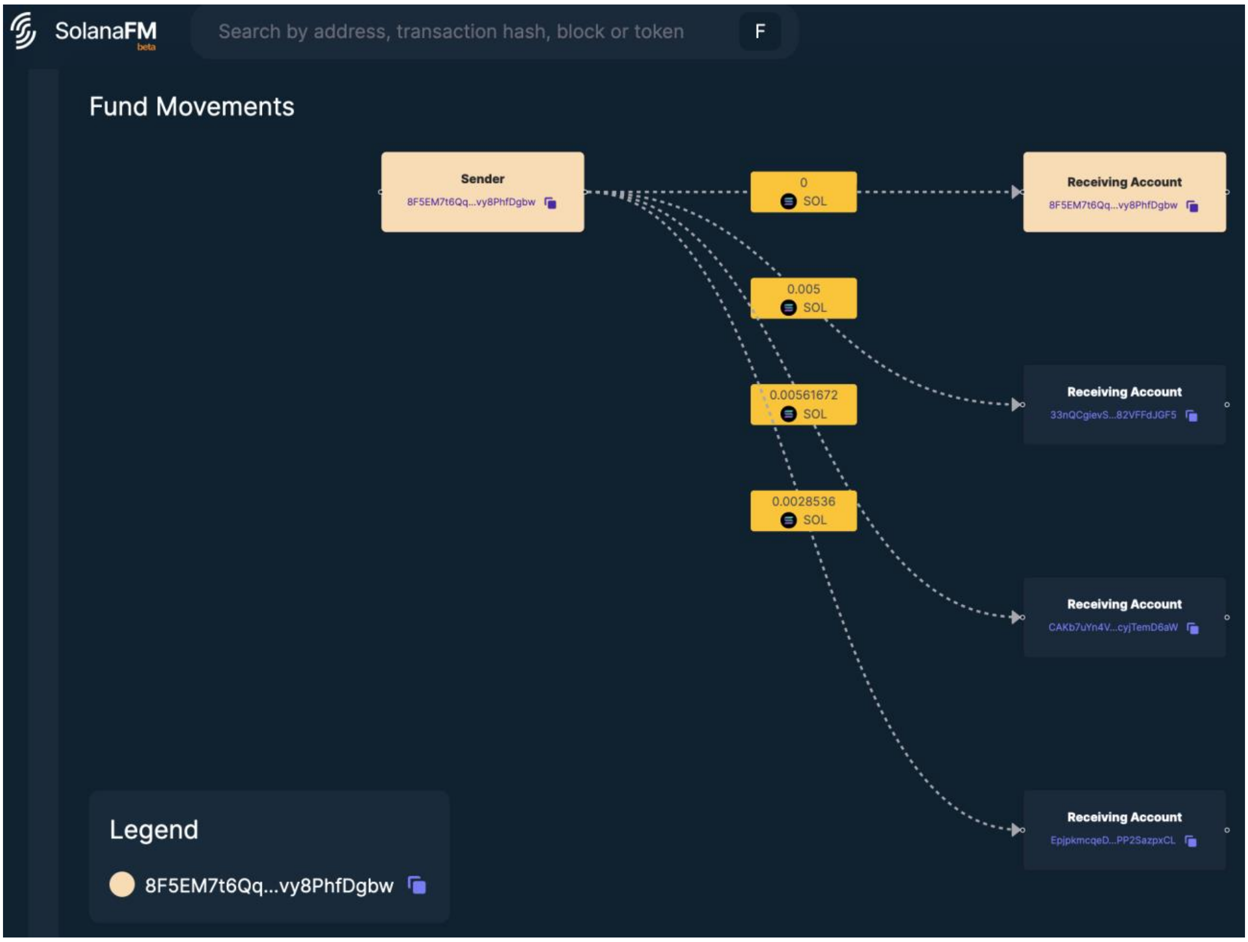


Figure 12

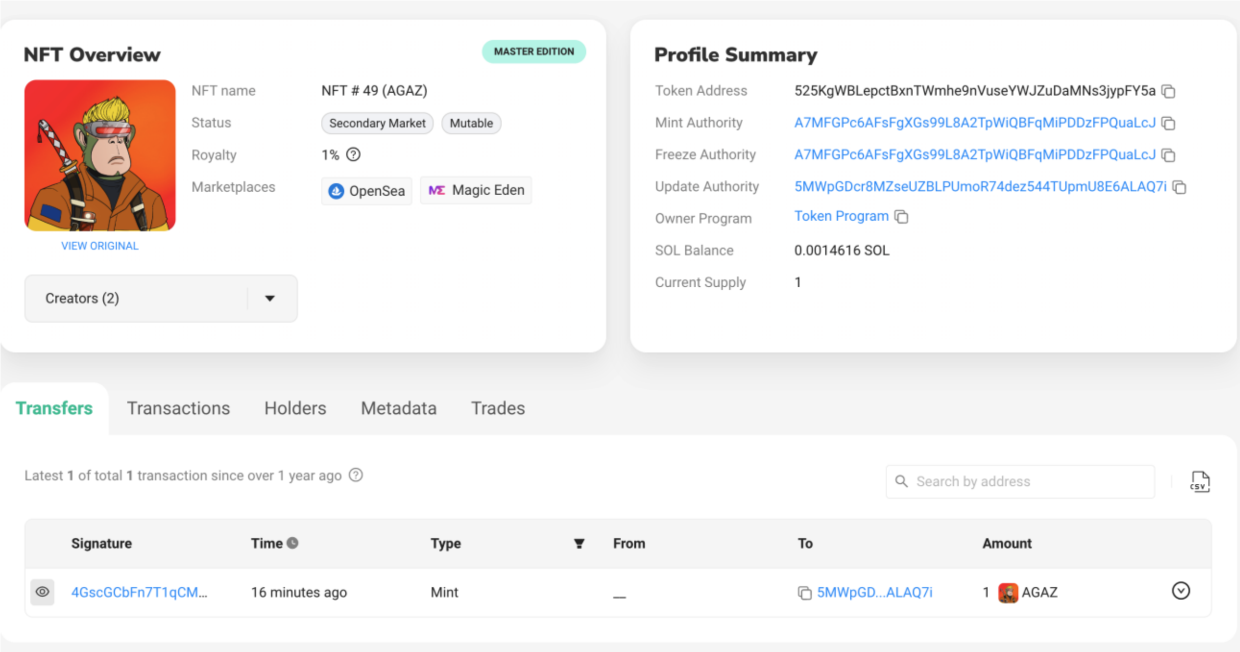


Figure 13

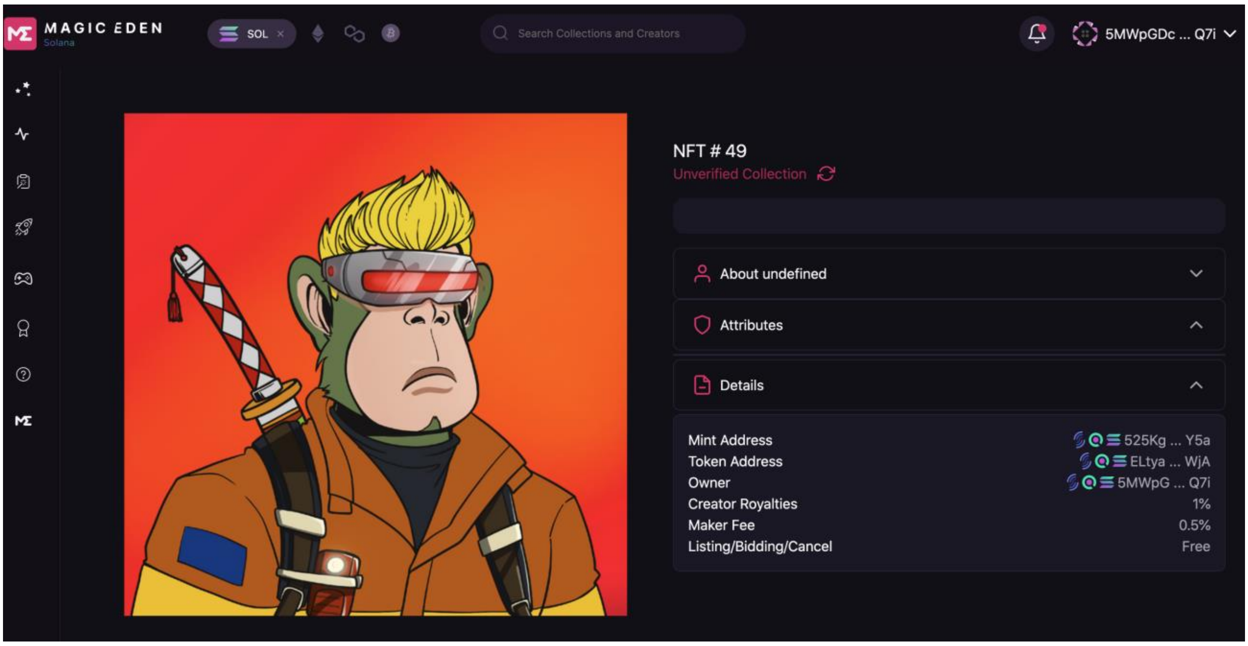


Figure 14

6. Conclusion and Future prospects

The project has delivered on its objectives to quantify the success of NFTs by analyzing different features of the assets and using machine learning models to predict their success. Firstly, the algorithm designed to deliver a visual representation/tree for different metrics of NFTs, such as sales volume, visual features, etc., is a powerful tool that can be used to analyze NFT data. The algorithm uses a decision tree algorithm to classify different NFTs based on their metrics, and the accuracy of the algorithm is evaluated using various metrics such as precision, recall, and F1-score. Additionally, the algorithm can be used to compare the statistical difference between two sets of means and standard deviations from two methods in a pair-wise manner using a t-test. Secondly, we have presented a methodology for using GNN to analyze NFT data and predict sales/profitability for future NFTs. Our results showed that our GNN algorithm outperformed a baseline linear regression model significantly. The GNN algorithm can be used to make more accurate predictions for NFT sales/profitability, which can be useful for NFT collectors and investors.

On these valuable outcomes for the two objectives, the analytical aspect of the NFT research is successful. Moreover, by demonstrating a simple and effective strategy to generate and launch NFTs into the market provide a holistic view from the process of collecting data of NFTs, analyzing their success all the way to launching your first NFT in the marketplace and looking forward to new advances in the space.

Although a comprehensive analysis of NFTs' features and metrics have been carried out to draw conclusions for determining growth and profitability of NFTs, future work in the NFT landscape presents various opportunities utilizing innovative techniques and perspectives.

6.1 Visuals extraction

The images or art of NFT hosted on marketplaces including OpenSea, are often not in a suitably high resolution whereby image classification can be carried on the features. Therefore, web scraping on popular series of NFTs can be carried out by navigating to the URL of their official landing pages/product pages. The visual features maybe extracted using an image classification deep convolutional neural network, PyTorch which is an implementation of AlexNet. The dataset of labelled images for feature extraction, in such a case, can be facilitated by ImageNet. The output of AlexNet is a 4096-value long vector is too dense for the predictive analysis for NFTs' sales growth. Therefore, Principle Component Analysis (PCA) helps to reduce dimensionality, thereby stripping the sample only to the main principle components, which is ideal for make data-driven predictions.

6.2 Account of unique applications of NFTs

A small account of diverse applications and technological implications that NFTs present in the field of autonomous vehicles and property rights, maybe discuss in future studies. All the features, systems and functionalities involved in those applications of NFTs may be demonstrated via simple, well-structured and easily comprehensible UML diagrams including Use Case, Sequence and Component digrams wherever appropriate and relevant. This contribution can help cement the multifaceted potential of NFTs beyond its traditional uses.

6.3 Negative externalities

One of the limitations of the models presented in this paper to predict future sales and success of NFTs, is that they do not take into account any negative externality which may have an unpredicited and indirect influence on the sales. Such factors can include geopolitical issues such as the Ukraine-Russia conflict, global coronavirus pandemic, country-wide economic recessions or even changes in market valuation due to mere comments/predictions from key economic players similar to ones we have seen influencers like Elon Musk make [10].

Correlating these factors of negative externality with the trend of NFTs' growth is beyond the scope of this project and provides a promising premise for future research

7. References

[1] Browne, R. (2022, March 10). *Trading in nfts spiked 21,000% to more than $17 billion in 2021, report says*. CNBC. Retrieved October 19, 2022, from https://www.cnbc.com/2022/03/10/trading-in-nfts-spiked-21000percent-to-top-17-billion-in-2021-report.html

[2] Yahoo! (n.d.). *Jefferies sees the NFT market reaching more than $80B in value by 2025*. Yahoo! Finance. Retrieved October 20, 2022, from https://finance.yahoo.com/news/jefferies-sees-nft-market-reaching-160604950.html

[3] *Understanding NFT smart contract*. Antier Solutions. (2022, August 1). Retrieved October 22, 2022, from https://www.antiersolutions.com/understanding-nft-smart-contract/

[4] FinSMEs. (2022, August 10). *The growing popularity of nfts*. FinSMEs. Retrieved October 19, 2022, from https://www.finsmes.com/2022/08/the-growing-popularity-of-nfts.html

[5] Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L. M., & Baronchelli, A. (2021, October 22). *Mapping the NFT revolution: Market trends, Trade Networks, and visual features*. Nature News. Retrieved October 21, 2022, from https://www.nature.com/articles/s41598-021-00053-8

[6] *NFT-based intelligence networking for connected and autonomous vehicles: A Quantum Reinforcement Learning Approach*. IEEE Xplore. (n.d.). Retrieved October 21, 2022, from https://ieeexplore.ieee.org/document/9852376/

[7] *NFTs and their uses beyond art*. The Economic Times. (n.d.). Retrieved October 21, 2022, from https://economictimes.indiatimes.com/tech/newsletters/ettech-unwrapped/nfts-and-their-uses-beyond-art/articleshow/87677153.cms?from=mdr

[8] *Non-Fungible Tokens Market Global Growth Drivers & Opportunities*. MarketsandMarkets. (n.d.). Retrieved February 6, 2023, from https://www.marketsandmarkets.com/Market-Reports/non-fungible-tokens-market-254783418.html

[9] NonFungible.com. (n.d.). *Q3 NFT market report 2022*. NonFungible.com. Retrieved February 6, 2023, from https://nonfungible.com/reports/2022/en/q3-quarterly-nft-market-report

[10] https://zipmex.com/learn/elon-musk-bitcoin-price/

[11] Vasan, K., Janosov, M., & Barabási, A.-L. (2022, February 17). *Quantifying NFT-driven networks in crypto art*. Nature News. Retrieved October 21, 2022, from https://www.nature.com/articles/s41598-022-05146-6

[12] Malwa, S. (2022, July 20). *Solana's New Daily Address Growth outpaces other blockchains, Sol Jumps*. CoinDesk Latest Headlines RSS. Retrieved October 21, 2022, from https://www.coindesk.com/tech/2022/07/20/solanas-new-daily-address-growth-outpaces-other-blockchains-sol-jumps/

[13] Levin, T. (n.d.). *NFTs could be the future of collecting - or a huge bubble. we talked to 3 experts about the risks to consider before buying in.* Business Insider. Retrieved October 21, 2022, from https://www.businessinsider.com/what-are-risks-of-investing-in-nft-2021-3

[14] Sharp, S. R. (2022, June 16). *NFT prices drop as "Crypto winter" descends*. Hyperallergic. Retrieved October 21, 2022, from https://hyperallergic.com/740883/nft-prices-drop-as-crypto-winter-descends/

[15] Henaff, M., Bruna, J., & LeCun, Y. (2015, June 16). *Deep convolutional networks on graph-structured data*. arXiv.org. Retrieved November 15, 2022, from https://arxiv.org/abs/1506.05163

[16] Duvenaud, D. K., Maclaurin, D., Iparraguirre, J., Bombarell, R., Hirzel, T., Aspuru-Guzik, A., & Adams, R. P. (1970, January 1). *Convolutional networks on graphs for learning molecular fingerprints*. Advances in Neural Information Processing Systems. Retrieved November 15, 2022, from https://papers.nips.cc/paper/2015/hash/f9be311e65d81a9ad8150a60844bb94c-Abstract.html

[17] Bruna, J., Zaremba, W., Szlam, A., & LeCun, Y. (2014, May 21). *Spectral networks and locally connected networks on graphs*. arXiv.org. Retrieved November 15, 2022, from https://arxiv.org/abs/1312.6203

[18] Kipf, T. N., & Welling, M. (2017, February 22). *Semi-supervised classification with graph Convolutional Networks*. arXiv.org. Retrieved November 15, 2022, from https://arxiv.org/abs/1609.02907

[19] *NFT-based intelligence networking for connected and autonomous vehicles: A Quantum Reinforcement Learning Approach*. IEEE Xplore. (n.d.). Retrieved November 15, 2022, from https://ieeexplore.ieee.org/document/9852376/

[20] Okonkwo, I. E. (2021, June 1). *NFT, copyright; and intellectual property commercialisation*. SSRN. Retrieved November 15, 2022, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3856154

[21] PTI / May 9, 2022. (n.d.). *Blocktickets, NFT and blockchain ticketing platform enters Indian market - Times of India*. The Times of India. Retrieved November 15, 2022, from https://timesofindia.indiatimes.com/blocktickets-nft-and-blockchain-ticketing-platform-enters-indian-market/articleshow/91440615.cms

[22] *The graph now supports Solana with substreams*. The Graph Blog. (n.d.). Retrieved November 15, 2022, from https://thegraph.com/blog/indexing-solana-substreams/

[23] *Moralis Web3 - Enterprise-grade Web3 apis*. Moralis Web3 | Enterprise-Grade Web3 APIs. (2022, November 15). Retrieved November 15, 2022, from https://moralis.io/

[23] Jain, S., Bruckmann, C., & McDougall, C. (2022, April 18). *NFT appraisal prediction: Utilizing search trends, public market data, linear regression and recurrent neural networks*. arXiv.org. Retrieved November 15, 2022, from https://arxiv.org/abs/2204.12932

Monthly Logs:

Monthly Log (March 2023)

Current Time: Mar 30, 2023 (Thu) 1:13:55 PM

Submission Deadline: Mar 31, 2023 (Fri) 11:59:59 PM

Date Last Edited: Mar 30, 2023 (Thu) 1:02 PM

Content✱:

Feb-Mar 2023:

An algorithm for creating a visual representation/tree for different NFT data, such as sales volume and visual attributes, is shown in the study. API calls gather data from open-source platforms, are preprocessed, and then analysed with a decision tree algorithm. The algorithm's precision is assessed using several metrics. The outcomes show how it may be used to pair-wise examine the statistical differences between two sets of averages and standard deviations from two approaches. The study describes the procedures for data preparation, model construction, prediction, and evaluation. The effectiveness of the decision tree model is graphically represented and evaluated using pertinent metrics

Monthly Log (January 2023)

Current Time: Feb 6, 2023 (Mon) 1:33:25 PM

Submission Deadline: Jan 31, 2023 (Tue) 11:59:59 PM

Date Last Edited: Feb 6, 2023 (Mon) 1:27 PM

Content✱:

1) Researched and selected current market trends for more valuable and recent insight into the NFT space to be included in the Introduction part of Interim Report II. 2) Planned the design requirements and gathered NFT artwork for series generation and launch purposes, by separating into layers for optimized use in Hash Lips Art engine node Java program. 3) Understood requirements about the specifications of NFTs to be deployed on the market, as well as the blockchain wallet and minting process allowing for first launch onto Mainnet. 4)More APIs are being researched extensively and machine learning models are being looked into in greater depth for processing data including visual components of NFTs and graphs of trade networks. Work underway and expected to span over the next couple of weeks.

✱ Required.

Save

Home (/student/index.do) / Monthly Logs (/student/log/index.do)

Monthly Log (December 2022)

Current Time: Feb 6, 2023 (Mon) 1:29:06 PM

Submission Deadline: Dec 31, 2022 (Sat) 11:59:59 PM

Date Last Edited: Dec 31, 2022 (Sat) 6:33 PM

Content✱:

B I U

- UML and workflow diagram prepared to show organization and structural overview of the project with various aspects.
- Extensive work on data collection and representation carried out and progressing smoothly
- NFT marketplace surveyed and NFT artist contacted to start designing artwork for market launch purposes

✱ Required.

Save

Home (/student/index.do) / Monthly Logs (/student/log/index.do)